

PAT-NO: JP02001256001A
DOCUMENT-IDENTIFIER: JP 2001256001 A
TITLE: DISK ARRAY DEVICE

PUBN-DATE: September 21, 2001

INVENTOR-INFORMATION:

NAME	COUNTRY
EZAKI, TOSHIHIRO	N/A
SOGO, YOSHIARI	N/A

ASSIGNEE-INFORMATION:

NAME	COUNTRY
MATSUSHITA ELECTRIC IND CO LTD	N/A

APPL-NO: JP2000070026
APPL-DATE: March 14, 2000

INT-CL (IPC): G06F003/06 , G06F013/10

ABSTRACT:

PROBLEM TO BE SOLVED: To solve a problem that an I/O processing function is remarkably lowered when a writing request to cause misalign is made.

SOLUTION: An Alignment judging processing part 23 judges whether or not a writing processing of data is Alignment by a logical address to indicate a starting position of writing of the received data and data size. When the Alignment judging processing part 23 judges the processing as Misalign, an Alignment state is set by supplementing invalid data so that the data size of the data becomes integer multiples of a parity calculation block when the data size is not the integer multiples of the parity calculation block by an invalid data supplement processing part 24 and converting the logical address so that the logical address coincides with offset when the logical address does not coincide with the offset of the parity calculation

block by an address conversion/ address management processing part
25.

COPYRIGHT: (C) 2001, JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-256001

(P2001-256001A)

(43) 公開日 平成13年9月21日 (2001.9.21)

(51) Int.Cl. ⁷	識別記号	F I	テ-マコード* (参考)
G 0 6 F 3/06	3 0 1	G 0 6 F 3/06	3 0 1 M 5 B 0 1 4
	5 4 0		5 4 0 5 B 0 6 5
13/10	3 4 0	13/10	3 4 0 B

審査請求 未請求 請求項の数 5 O L (全 19 頁)

(21) 出願番号 特願2000-70026 (P2000-70026)

(22) 出願日 平成12年3月14日 (2000.3.14)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 江崎 俊裕

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 十河 美存

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(74) 代理人 100097445

弁理士 岩橋 文雄 (外2名)

Fターム(参考) 5B014 EB05

5B065 BA01 CA07 CA30 CC08 CE04

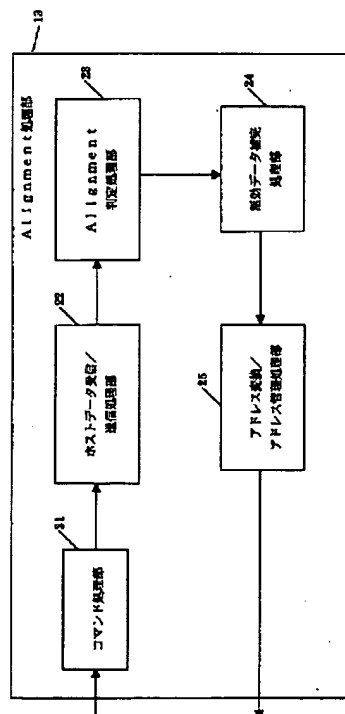
CS05

(54) 【発明の名称】 ディスクアレイ装置

(57) 【要約】

【課題】 Misalignが発生する書き込み要求があった場合、I/O処理機能が著しく低下する。

【解決手段】 Alignment判定処理部23は受信したデータの書き込み開始位置を示す論理アドレスとデータサイズよりデータの書き込み処理がAlignmentであるかどうかを判定する。Alignment判定処理部23がMisalignと判定した場合、無効データ補完処理部24はデータのデータサイズがパリティ計算ブロックの整数倍でない場合は、データサイズがパリティ計算ブロックの整数倍になるよう無効データを補完し、アドレス変換/アドレス管理処理部25は論理アドレスがパリティ計算ブロックのオフセットと一致しない場合は、論理アドレスがオフセットと一致するよう論理アドレスを変換することによりAlignment状態にする。



【特許請求の範囲】

【請求項1】 複数のディスク装置を使用してホストからのデータの書き込みおよび読み込みの制御を行うディスクアレイ装置であって、

前記ホストからデータの書き込み開始アドレスとデータとからなるデータの書き込み要求を受け、前記データをあらかじめ定められたデータサイズのデータブロックに分割して複数のディスク装置に書き込む場合、

前記書き込み開始アドレスが前記データの書き込み対象となるデータブロックのオフセット値を示すブロック番号と一致するかどうかを判定し、一致していない場合は前記データの書き込み開始アドレスを前記データの書き込み対象となるデータブロックの次に位置するデータブロックのオフセット値を示すブロック番号に一致するようシフトして、前記書き込み開始アドレスを更新し、前記書き込み開始アドレスの更新情報を記憶するアラインメント処理手段と、

前記アラインメント処理手段により更新された書き込み開始アドレスより前記データを前記データブロックのデータサイズ毎に分割して前記複数のディスク装置へ書き込むよう制御するディスクアレイコントローラとを有するディスクアレイ装置。

【請求項2】 複数のディスク装置を使用してホストからのデータの書き込みおよび読み込みの制御を行うディスクアレイ装置であって、

前記ホストからデータの書き込み開始アドレスとデータとからなる書き込み要求を受け、前記データをあらかじめ定められたデータサイズのデータブロックに分割して複数のディスク装置に書き込む場合、

前記データのデータサイズが前記データブロックのデータサイズの整数倍であるかどうかを判定し、整数倍になっていない場合は、前記データのデータサイズが前記データブロックのデータサイズの整数倍になるように前記データに前記データと独立した無効データを付加するアラインメント処理手段と、

前記アラインメント処理手段により無効データを付加されたデータを前記書き込み開始アドレスより前記データブロックのデータサイズ毎に分割して前記複数のディスク装置へ書き込むよう制御するディスクコントローラとを有するディスクアレイ装置。

【請求項3】 複数のディスク装置を使用してホストからのデータの書き込みおよび読み込みの制御を行うディスクアレイ装置であって、

前記ホストからデータの書き込み開始アドレスとデータとからなる書き込み要求を受け、前記データをあらかじめ定められたデータサイズのデータブロックに分割して複数のディスク装置に書き込む場合、

前記書き込み開始アドレスを前記書き込み開始アドレスと前記データの書き込み対象となるデータブロックのオフセット値を示すブロック番号と一致するようあらかじめ

め設定されているシフト量だけシフトして前記書き込み開始アドレスを更新するアラインメント処理手段と、前記アラインメント処理手段により更新された書き込み開始アドレスより前記データを前記データブロックのデータサイズ毎に分割して前記複数のディスク装置へ書き込むよう制御するディスクアレイコントローラとを有するディスクアレイ装置。

【請求項4】 アラインメント処理手段の機能を備えたディスクアレイコントローラを有する請求項1ないし3のディスクアレイ装置。

【請求項5】 ディスクアレイコントローラはRAIDコントローラであることを特徴とする請求項4記載のディスクアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数のディスク装置を並列的に動作させてデータの読み込み、および書き込み制御を行うディスクアレイ装置に関する。

【0002】

【従来の技術】従来の技術として、例えば、特開平9-305328号公報に示されている技術がある。以下、図面に基づいて従来例を説明する。

【0003】図13に従来のディスクアレイ装置のブロック図を示す。

【0004】ホスト1301から見るとディスクアレイコントローラ1303以降のブロックはホストアダプタ1302を介して1台のディスク装置として見える。ホスト1301はホストアダプタ1302を介してデータをディスクアレイコントローラ1303へ送信する。この場合、ディスクアレイコントローラ1303は、受信したデータを4分割し、各データをデバイスコントローラ1304a~1304dに送信すると共に、分割データに対するパリティデータを生成しデバイスコントローラ1304eへ送信する。デバイスコントローラ1304a~1304dは受信した分割データを各々データ格納用ディスク装置1305a~1305dに記録し、デバイスコントローラ24eは受信したパリティデータをパリティデータ格納用ディスク装置25eに記録する。

【0005】このように従来のディスクアレイ装置は、複数のディスク装置（ハードディスク装置）を使用することにより、高速に大量のデータにアクセスし、ディスク故障時におけるデータの冗長性を実現することができ、単独のディスク装置よりも高い信頼性と性能を実現している。これは、1987年に米国のカリフォルニア大学バークレイ校のデビッド・A. パターソン(David A. Patterson)教授らが提唱したRAID (Redundant Arrays of Inexpensive Disks) と呼ばれるものである。すなわち、RAIDはデビッド・A. パターソン教授らの論文に由来する呼び方である。

【0006】ディスクアレイ方式RAIDは、1から5

までのレベル（以下、レベル1から順にRAID1、RAID2、RAID3、RAID4、RAID5と記す）に分類されており、RAID3からRAID5は、ディスク装置の故障時にデータを回復させるためのパリティデータを保持する。例えば、RAID3は、入力データを分割（ストライピング）し、インターリーブをかけて分割したデータを複数のディスク装置に分散して格納するものである。この場合、分割したデータに対するエラー訂正符号を格納するディスク装置を冗長ディスク装置とし、冗長ディスク装置のデータをパリティデータとしたものである。

【0007】以下、RAID3におけるデータ書き込み処理の流れを図14、図15を用いて説明する。図14はディスクアレイコントローラ1303の詳細なブロック図であり、図15はディスクアレイコントローラ1303の処理の流れを示すフローチャートである。

【0008】まず、書き込み処理について説明する。

【0009】コマンド処理部1402はホスト1301からの書き込み要求を受信し（図15のステップ1501）、ホスト1301からの要求が書き込み要求であることを判定し（図15のステップ1502）、データの書き込み処理に必要な処理を行う。コマンド処理部1402は、データの書き込み開始位置を示す論理アドレスとデータの送信をホスト1301に要求する。ホスト1301は、コマンド処理部1402の要求を受け取ると、直ちに論理アドレスとデータをコマンド処理部1402に送信する。コマンド処理部1402は、ホスト1301より送信された書き込み要求、論理アドレス、およびデータをホストデータ受信/送信処理部1403を介してAlignment判定処理部1404に送信する。

【0010】Alignment判定処理部1404は、受信した論理アドレスとデータのデータサイズから、データ書き込み時の処理が、以下の（1）、（2）の条件を満たすかどうかを判定する。（図15のステップ1503）。

（1）データサイズブロックが、ハードディスクの最小単位1セクタ（512Byte）の整数倍であり、かつパリティ計算ブロック（512Byteの整数倍サイズ）の整数倍である。

（2）データの書き込み開始位置を示すアドレスが、パリティ計算ブロックのAlignment（アラインメント）境界であるオフセットと一致する。

【0011】Alignment判定処理部1404は、上記2つの条件を満たす状態をAlignment、いずれか一つでも満たさない状態をMisalign（ミサライン）と定義する。Alignment判定処理部1404は、Alignment判定を行った後、書き込み要求、論理アドレス、データ、および判定結果をホストデータ分割/結合処理部1405に送信する。

る。

【0012】以下、Alignment判定処理部1404以降の処理を4KB（1KB=1024Byte）のパリティ計算ブロックに、4KBの固定長のデータを書き込む場合を例にとって図16を用いて説明する。図16において、パリティ計算ブロックは4KBであり、パリティ計算ブロックは8個のデータサイズブロックで構成され、1つのデータサイズブロックは512B（Byte）を示している。また、矢印は書き込み開始位置を示している。

【0013】まず、判定結果がAlignmentである場合の処理を説明する。

【0014】図16の（a）にAlignmentデータの一例を示す。Alignment判定処理部1404は図16の（a）のデータはデータサイズ4KB（=512×8B）であることから条件（1）を満たし、書き込み開始位置がパリティ計算ブロックのAlignment境界であるオフセットと一致していることから条件（2）も満たしていると判断し、Alignmentであると判定する。Alignment判定処理部1404は判定結果、書き込み要求、論理アドレス、データをホストデータ分割/結合処理部1405へ送信する。

【0015】ホストデータ分割/結合処理部1405は受信した論理アドレスと判定結果をアドレス指定処理部1406に送信すると同時に、受信したデータを分割してアドレス指定処理部1406に送信する（図15のステップ1504）。この場合、4台のディスク装置が接続されているので、データを4分割する。アドレス指定処理部1406は受信した判定結果に従って論理アドレスのアドレス変換を行う。この場合、判定結果がAlignmentであることから、アドレス指定処理部1406は受信した論理アドレスに対するアドレス変換の処理を行わず、パリティデータに対してのみアドレスを割り当て、受信した判定結果、書き込み要求、論理アドレス、パリティデータのアドレス、データをパリティ処理部1407へ送信する（図15のステップ1505）。

ここで、この時点ではパリティデータは存在しないが、パリティデータサイズは固定サイズであるため、先にアドレスだけ割り当てている。パリティ処理部1407は、受信した分割データよりパリティデータを作成し、分割データ、パリティデータ、書き込み要求、論理アドレス、パリティデータのアドレスをディスク装置1305に送信する（図15のステップ1506）。ディスク装置1305は、受信した分割データとパリティデータを指定されたアドレスに従って記録する。

【0016】次に、判定結果がMisalignである場合の処理を説明する。この場合、連続する2つの4KBのデータ、データ1、データ2を順次書き込む場合を例にとって説明する。

【0017】まず、データ1の書き込み処理を説明す

る。

【0018】図16の(b)の塗りつぶし部分にデータ1を示す。同図に示すようにデータ1は4KBのデータであり、矢印に示す書き込み開始位置は[0001h]である。アドレス[0001h]から4KBのデータを書き込むと、図16の(b)に示すようにブロック[01h]とブロック[02h]に跨って記録される。

【0019】Alignment判定処理部1404は、データ1はデータサイズが4KBであることから条件(1)を満たし、書き込み開始位置[0001h]が図16の(b)の矢印位置にあることから、パリティ計算ブロックのAlignment境界であるオフセットと一致しないため、Alignmentの条件(2)を満たしていないと判断する。すなわち、Alignment判定処理部1404はデータ1の書き込み処理はMisalignであると判定し、判定結果、書き込み要求、論理アドレス[0001h]、およびデータ1をホストデータ分割/結合処理部1405に送信する(図15のステップ1503)。

【0020】ホストデータ分割/結合処理部1405は、受信した論理アドレス[0001h]、判定結果、および、書き込み要求をアドレス指定処理部1406に送信すると同時に、データ1の分割処理を実行した後、分割データをアドレス指定処理部1406に送信する(図15のステップ1507)。アドレス指定処理部1406は、受信した論理アドレス[0001h]から、分割データとパリティデータを実際の記録先を示すアドレス(実アドレス)の指定を行い、論理アドレス、実アドレスをアドレス管理テーブルに保有する(図15のステップ1508)。(表1)にアドレス指定処理部1406が保有するアドレス管理テーブルを示す。

【0021】

【表1】

論理アドレス (ホストからのアドレス)	実アドレス (ディスクに書き込むブロック番号)
0001h	01h

【0022】(表1)に示すように、アドレス管理テーブルはデータをホスト1301から送信された論理アドレスと実際にディスクに書き込むブロック番号を示す実アドレスで構成されている。(表1)に示すように、データ1の論理アドレスは[0001h]、実アドレス、すなわちブロック番号は[01h]である。

【0023】アドレス指定処理部1406は、書き込み要求、実アドレス[01h]、分割データ、判定結果をパリティ処理部1407に送信する。パリティ処理部1407は、分割データよりパリティデータを作成し、分割

データ、パリティデータ、書き込み要求、および実アドレス[01h]をディスク装置1305に送信する(図15のステップ1509、ステップ1510)。この場合、データ1は最初の書き込みであり、書き込み対象のパリティ計算ブロックにおいてパリティデータは存在しないため、分割データのみでパリティデータを生成する。ディスク装置1305は、受信した書き込み要求、実アドレスに従って、分割データとパリティデータを記録する。

【0024】次にデータ2の書き込み処理について説明する。

【0025】図16の(c)の斜線部にデータ2を示す。図16の(c)に示すように、データ2は4KBのデータであり、書き込み開始位置アドレスを示す論理アドレスは矢印が示す[0009h]であり、ブロック[02h]とブロック[03h]に跨って記録される。

【0026】Alignment判定処理部1404は、データ2はデータサイズが4KBであることから条件(1)を満たし、図6の(c)に示す書き込み開始位置[0009h]がパリティ計算ブロックのAlignment境界であるオフセットと一致しないため、Alignmentの条件(2)を満たしていないと判断する。従って、Alignment判定処理部1404はMisalignであると判定し、判定結果、書き込み要求、論理アドレス[0009h]、および、データ2をホストデータ分割/結合処理部1405に送信する(図15のステップ1503)。

【0027】ホストデータ分割/結合処理部1405は、受信した論理アドレス[0009h]、判定結果、および書き込み要求をアドレス指定処理部1406に送信すると同時に、データ2の分割処理を実行した後、分割データをアドレス指定処理部1406に送信する(図5のステップ1507)。アドレス指定処理部1406は、保有する(表1)に示すアドレス管理テーブルとデータ2の論理アドレス[0009h]から、分割データとパリティデータの実アドレス指定を行い、アドレス情報をアドレス管理テーブルに格納する(図5のステップ1508)。(表2)にアドレス指定処理部1406が更新したアドレス管理テーブルを示す。

40 【0028】

【表2】

論理アドレス (ホストからのアドレス)	実アドレス (ディスクに書き込むブロック番号)
0001h	01h
0009h	02h

【0029】(表2)に示すように、データ1のアドレス情報の次に、データ2のアドレス情報である、論理ア

ドレス[0009h]と、データ2の書き込み開始位置があるブロックのブロック番号[02h]が実アドレスとして格納されている。

【0030】アドレス指定処理部1406は、判定結果、実アドレス[02h]、書き込み要求、及び分割データをパリティ処理部1407に送信する。パリティ処理部1407は、受信した分割データ、判定結果、および実アドレス[02h]より、データ2のパリティデータを計算する。まず、図6の(b)のパリティ計算ブロック[0001h]に記録されているデータ1の一部であるデータAとパリティ計算ブロック[02h]のパリティデータの読み出しを行う(図5のステップ1509)。次に、パリティ処理部1407は、読み出したデータAとパリティ計算ブロック[02h]に格納されるデータ2の分割データよりパリティデータを作成する(図5のステップ1510)。パリティ処理部1407は、データA、パリティ計算ブロック[02h]に格納されるデータ2の分割データ、新たに生成されたパリティデータ、および、実アドレス[02h]を、ディスク装置1305に送信する。ディスク装置1305は、受信した実アドレスに従って、受信したデータ2の分割データ、データA、および、パリティデータを記録する。

【0031】また、パリティ処理部1407はパリティ計算ブロック[03h]に記録されるデータ2の残りの分割データとデータが格納されていない7データサイズブロックに存在するデータに対してパリティデータを生成し、書き込み要求、データ2の残りの分割データ、生成されたパリティデータ、実アドレス[03h]をディスク装置1305に送信する。ディスク装置1305は受信したデータ2、パリティデータを実アドレスに従って記録する。

【0032】この様に、従来のディスクアレイ装置では、データの書き込み処理がMisalign、かつ、データを書き込むパリティ計算ブロックに既に記録されているデータがある場合は、そのデータとデータに対するパリティデータの読み出しを行ってから、再度パリティデータの生成を行っている。

【0033】

【本発明が解決しようとする課題】しかしながら、上記の従来のディスクアレイ装置は、I/O処理がAlignmentの時に最高の性能を発揮するディスクアレイ装置に対して、固定データサイズのデータの連続I/O処理を行う際に、固定データサイズがパリティ計算ブロックサイズの整数倍ではない場合や、データの書き込み開始位置を示すアドレスがパリティ計算ブロックのオフセットと一致しない場合にMisalignが発生するため、パリティデータ生成時にAlignmentの時にはないデータの読み出し処理が必要となり、I/O処理性能が著しく低下するという問題が生じる。

【0034】本発明は、上記従来の問題点を解決するも

ので、Misalignが発生するI/O処理に対して、後続のI/O処理がAlignmentで行えるようにするディスクアレイ装置を提供する。

【0035】

【課題を解決するための手段】この目的を達成するために本発明のディスクアレイ装置は、Misalignが発生する書き込み要求があった場合、後続のデータの書き込みがAlignmentになるようデータの補完やアドレス変換を行うAlignment処理機能を有している。

【0036】この構成によって、本発明のディスクアレイ装置は、Misalignが発生する書き込み要求があった場合、後続のデータの書き込みをAlignmentで行うことができる。

【0037】

【実施の形態】本発明の第1の発明は、あらかじめ定められたデータサイズを処理単位として、ホストからの要求に従ってデータ処理を行うディスクアレイ装置であって、前記ホストから受信した書き込み開始位置を前記処理単位のアラインメント境界と一致するよう変換し、変換情報を記憶して管理するアラインメント処理手段を有するディスクアレイ装置、書き込み開始位置をパリティ計算ブロックのアラインメント境界に一致させることにより、後続のデータのI/O処理をアラインメントで行うことができる。

【0038】本発明の第2の発明は、あらかじめ定められたデータサイズを処理単位として、ホストからの要求に従ってデータ処理を行うディスクアレイ装置であって、前記ホストから受信したデータのデータサイズが前記処理単位のデータサイズの整数倍になるように前記データと独立した無効データを付加するアラインメント処理手段を有するディスクアレイ装置であって、データサイズをパリティ計算ブロックの整数倍にすることによって、後続のデータのI/O処理をアラインメントで行うことができる。

【0039】本発明の第3の発明は、あらかじめ定められたデータサイズを処理単位として、ホストからの要求に従ってデータ処理を行うディスクアレイ装置であって、前記ホストから受信した書き込み開始位置を前記処理単位のアラインメント境界と一致するようあらかじめ定められたシフト値だけシフトするアラインメント処理手段を有するディスクアレイ装置であって、シフト値を誤り訂正用のパリティ計算ブロックのアラインメント境界に一致させる値に設定することにより、後続のデータのI/O処理をアラインメントで行うことができる。

【0040】以下に、本発明によるディスクアレイ装置について、添付図面を参照しながら説明する。

【0041】(実施の形態1) 実施の形態1では、Alignment処理機能をディスクアレイコントローラ

の外部に持たせ、Misalignが発生する書き込み要求をAlignmentの書き込み要求にする場合について説明する。

【0042】図1は、本発明の一実施の形態のディスクアレイ装置の構成を示すブロック図である。

【0043】図1において、11はI/O処理を指示するホスト、12はホスト11と接続する装置を制御するホストアダプタ、13はホストアダプタ12より入力されるデータにAlignment（アラインメント）処理を行うAlignment処理部、14は書き込み時にはAlignment処理部13より入力されたデータを分割し、分割したデータのパリティデータを生成し、読み込み時にはデータのパリティチェックと分割データを統合して1つのデータとするディスクアレイコントローラ、15a～15eはディスクアレイコントローラ14より入力されたデータのI/O処理を制御するデバイスコントローラ、16a～16eはそれぞれデバイスコントローラ15a～15eの制御に従ってデータのI/O処理を行うディスク装置である。

【0044】以上のように構成されたディスクアレイ装置について、図面を参照しながら説明する。

【0045】ホスト11は、ホストアダプタ12経由で、書き込み要求とデータの書き込み開始位置（以下、論理アドレスと記す）、およびデータをAlignment判定処理部13に送信する。Alignment判定処理部13は、ホスト11からの書き込み要求を受信し、受信したデータの論理アドレスとデータから、データの書き込み処理がAlignmentであるかどうかの判定を行う。

【0046】以下、Alignment処理部13の詳細な動作を図2、図3を用いて説明する。図2はAlignment処理部13の詳細な構成を示すブロック図、図3はAlignment処理部13の処理を示すフローチャートである。

【0047】図2において、21はホスト11から受信したコマンドを処理するコマンド処理部、22はコマンド処理部21から出力されたデータを受信し、受信したデータをデータの書き込み処理に対してAlignment判定を行うAlignment判定処理部23へ送信するホストデータ受信/送信処理部、24はAlignment判定処理部23より出力された判定結果に応じて、Alignment判定処理部23より出力された無効データを補完する無効データ補完処理部、25はAlignment判定処理部23より出力された判定結果に応じて無効データ補完処理部24より出力されたデータの論理アドレスを実アドレスへ変換すると共に、双方のアドレスの管理を行うアドレス変換/アドレス管理処理部である。

【0048】コマンド処理部21は、ホスト11から送信されてきたアクセス要求が、書き込み処理であるか読

み込み処理であるかの判定を行う。書き込み処理の場合は、手順通り残りの4つの処理ブロックでそれぞれの処理が行われ、読み込み処理の場合4つの処理ブロックでは処理は行われず、ディスクアレイコントローラ14にホスト11より受信したデータが送信される。（図3のステップ301～ステップ302）。

【0049】以下、ホスト11からの要求が書き込み処理要求である場合を説明する。

【0050】コマンド処理部21はホスト11より入力された書き込み要求、論理アドレス、およびデータをホストデータ受信/送信処理部22を介してAlignment判定処理部23に送信する。

【0051】Alignment判定処理部23は、受信したデータを無効データ補完処理部24に送信すると同時に、受信した論理アドレスとデータのデータサイズよりデータの書き込み処理がAlignmentであるかどうかの判定を行い、判定結果を無効データ補完処理部24に送信する（図3のステップ303）。

【0052】ここで、Alignmentとは、以下の（1）（2）の条件を満たす状態を指す。

（1）データサイズブロックが、ハードディスクの最小単位1セクタ（512B）の整数倍であり、かつパリティ計算ブロック（512Bの整数倍サイズ）の整数倍である。

（2）データの書き込み開始位置を示すアドレスが、パリティ計算ブロックのAlignment境界であるオフセットと一致する。

【0053】また、2つの条件のいずれか一つでも満たさない状態をMisalign（ミスアライン）と定義する。

【0054】Alignment判定処理部23の判定結果として、Alignmentであると判定された場合は、残り2つの処理ブロックでは処理が施されず、受信したデータはそのままディスクアレイコントローラ14へ送信される。Misalignと判定された場合は、Misalignの状態に応じて、残りの2つの処理ブロックで処理が施される。

【0055】無効データ補完処理部24は、Alignment判定処理部23から送信されてきた判定結果より、データが条件（1）を満たしている場合は、直ちに、書き込み要求、データ、論理アドレス、および判定結果を、アドレス変換/アドレス管理処理部25に送信する。また、データが条件（1）を満たしていない場合、すなわち、Misalign状態の場合、無効データ補完処理部24は、データに無効データを補完して、パリティ計算ブロックの整数倍になるように処理し、アクセス要求、無効データが追加された新たなデータ、論理アドレス、および判定結果をアドレス変換/アドレス管理処理部25に送信する（図3のステップ304～ステップ305）。

11

【0056】アドレス変換/アドレス管理処理部25は、受信した判定結果より、受信したデータが条件(2)を満たしている場合、直ちに、書き込み要求、データ、論理アドレス、およびデータをディスクアレイコントローラ14に送信すると同時に、論理アドレスより、保有していたアドレス管理テーブルの更新を行う。条件(2)を満たしていない場合、データの書き込み開始位置を書き込もうとしていたパリティ計算ブロックの次のパリティ計算ブロックのオフセットに変更する実アドレスの割り当てを行う。また、アドレス変換/アドレス管理処理部25は、論理アドレスと仮想アドレスの対応を示すアドレス情報を管理しているアドレス管理テーブルを保有している(図3のステップ306～ステップ309)。

【0057】以下、Misalign状態の3つの場合におけるAlignment判定処理部23、無効データ補完処理部24、アドレス変換/アドレス管理処理部25の処理を詳細に説明する。

【0058】まず、条件(1)のみを満たさないMisalign状態の場合について説明する。

【0059】この場合のデータの一例を図4の(a)の塗りつぶし部分に示す。図4において、処理単位となるパリティ計算ブロックサイズは4KB、パリティ計算ブロックにおける個々のデータブロックサイズは512Bである。図4の(a)のデータは、論理アドレス[0000h]、データサイズが6KBのデータである。

【0060】Alignment判定処理部23は、受信したデータを直ちに無効データ補完処理部24に送信すると同時に、受信した論理アドレスとデータサイズより、データの書き込み処理がAlignmentであるかどうかを判定する。Alignment判定処理部23は、論理アドレス[0000h]はパリティ計算ブロックのオフセット値を示すブロック番号[00h]と一致しているが、データサイズ6KBはパリティ計算ブロックサイズの4KBの倍数ではないことから、条件(1)のみを満たさないMisalignという判定結果Aを、無効データ補完処理部24に送信する。

【0061】無効データ補完処理部24は、Alignment判定処理部23から送信されてきた判定結果Aより、条件(1)を満たさないMisalignであると判断し、図4の(a)に示す6KBのデータがパリティ計算ブロック4KBの最小公倍数の8KBになるように2KBの無効データを補完し、書き込み要求、無効データが追加された8KBのデータ、論理アドレス[0000h]、そして判定結果Aをアドレス変換/アドレス管理処理部24に送信する。図4(b)に無効データを追加された8KBのデータを示す。同図に示すように、図4の(b)のデータは図4の(a)のデータに斜線部に示す2KBの無効データを補完することで、パリティ計算ブロックの整数倍(2倍)のデータとなっている。

12

【0062】アドレス変換/アドレス管理処理部25は、受信した判定結果Aより、条件(2)を満たしていると判断し、直ちに、書き込み要求、8KBのデータ、論理アドレス[0000h]をディスクアレイコントローラ14に送信すると同時に、論理アドレス[0000h]より、実アドレスの割り当てを行い、保有していたアドレス管理テーブルの更新を行う。(表3)にアドレス変換/アドレス管理処理部25が保有するアドレス管理テーブルを示す。

【0063】

【表3】

論理アドレス (ホストからのアドレス)	実アドレス (変換アドレス)
0000h	0000h
000Ch	0010h

【0064】(表3)に示すように、アドレス管理テーブルはデータの論理アドレスと実アドレスの対応した形で構成される。(表3)に示すように、判定結果Aより受信したデータは条件(2)を満たしているため、論理アドレス[0000h]と実アドレス[00h]は一致している。

【0065】次に、条件(2)のみを満たさないMisalignの場合について説明する。

【0066】この場合のデータの一例を図5の(a)の塗りつぶし部分に示す。図5において処理単位となるパリティ計算ブロックサイズは4KB、パリティ計算ブロックにおける個々のデータブロックサイズは512Bである。図5の(a)のデータは、データサイズが4KB、論理アドレスが[0001h]のデータである。

【0067】Alignment判定処理部23は受信したデータを直ちに、無効データ補完処理部24に送信すると同時に、受信した論理アドレスとデータのデータサイズより、データの書き込み処理がAlignmentであるかどうかを判定する。Alignment判定処理部23は、データサイズ4KBはパリティ計算ブロックサイズが4KBの整数倍であるが、論理アドレス[0001h]とパリティ計算ブロックのオフセット値を示すブロック番号[00h]とが一致していないことから、条件(2)のみを満たさないMisalign状態という判定結果Bを、無効データ補完処理部24に送信する。

【0068】無効データ補完処理部24は、Alignment判定処理部23から送信された判定結果Bより、受信したデータは条件(1)を満たしているため補完処理をする必要がないと判断し、送信されてきた書き込み要求、データ、論理アドレス[0001h]、および判定結果Bをアドレス変換/アドレス管理処理部24に

送信する。

【0069】アドレス変換／アドレス管理処理部25は、受信した判定結果Bより、受信したデータが条件(2)を満たしていないと判断し、受信した論理アドレスに対してアドレス変換を行う。アドレス変換／アドレス管理処理部25は、受信したデータの論理アドレス[0001h]を、書き込みを行おうとしていたパリティ計算ブロック[00h]の次のパリティ計算ブロックのオフセット値を示すブロック番号[01h]にシフトさせる為に、論理アドレス[0001h]に対して、7データ

サイズブロックだけシフトした実アドレス[0008h]の割り当てを行い、その変換情報をアドレス管理テーブルに保有する。

【0070】(表4)にアドレス変換／アドレス管理処理部25が保有しているアドレス管理テーブルを示す。

【0071】

【表4】

論理アドレス (ホストからのアドレス)	実アドレス (変換アドレス)
0001h	0008h
0009h	0010h

【0072】(表4)に示すように、データ1の論理アドレス[0001h]に対応する実アドレスとして[0008h]が割り当てられているという変換情報が記述されている。

【0073】アドレス変換／アドレス管理処理部25はアドレス管理テーブルの更新完了後、書き込み要求、データ、実アドレス[0008h]の情報をディスクアレイコントローラ14に送信する。

【0074】最後に、条件(1)、(2)を満たさないMisalign状態の場合について説明する。

【0075】この場合のデータの一例を図6の(a)の塗りつぶし部分に示す。図6において、処理単位となるパリティ計算ブロックサイズは4KB、処理単位となるパリティ計算ブロックにおける個々のブロックサイズは512Bである。図6の(a)のデータは、論理アドレス[0001h]、データサイズが6KBのデータである。

【0076】Alignment判定処理部23は受信したデータを直ちに、無効データ補完処理部24に送信すると同時に、受信した論理アドレス[0001h]とデータのデータサイズより、データの書き込み処理がAlignmentであるかどうかを判定する。Alignment判定処理部23は、データサイズ6KBがパリティ計算ブロックサイズの4KBの倍数ではなく、論理アドレス[0001h]とパリティ計算ブロックのオフセット値を示すブロック番号[00h]と一致していないこ

とから、条件(1)、(2)を満たさないMisalignであるという判定結果Cを無効データ補完処理部24に送信する。

【0077】無効データ補完処理部24は、Alignment判定処理部23から送信されてきた判定結果Cより、条件(1)を満たさないMisalign状態であると判断し、図6の(a)に示す6KBのデータがパリティ計算ブロック4KBとの最小公倍数8KBになるように2KBの無効データを補完し、書き込み要求、無効データが追加された新たな8KBデータ、論理アドレス[0001h]、および判定結果Cをアドレス変換／アドレス管理処理部25に送信する。図6の(b)に無効データを追加された8KBのデータを示す。同図に示すように、図6の(b)のデータは図6の(a)のデータに斜線部に示す2KBの無効データを補完することで、パリティ計算ブロックの整数倍(2倍)のデータとなっている。

【0078】アドレス変換／アドレス管理処理部25は、受信した判定結果Cより、受信したデータが条件(2)を満たしていないと判断し、受信した論理アドレスに対してアドレス変換を行う。アドレス変換／アドレス管理処理部25は、受信したデータの論理アドレス[0001h]を、書き込みを行おうとしていたパリティ計算ブロック[00h]の次のパリティ計算ブロックのオフセット値を示すブロック番号[01h]にシフトさせる為に、論理アドレス[0001h]に対して、7データサイズブロックだけシフトした実アドレス[0008h]の割り当てを行い、その変換情報をアドレス管理テーブルに保有する。

【0079】(表5)にアドレス変換／アドレス管理処理部25が保有しているアドレス管理テーブルを示す。

【0080】

【表5】

論理アドレス (ホストからのアドレス)	実アドレス (変換アドレス)
0001h	0008h
000Dh	0018h

【0081】(表5)に示すように、論理アドレス[0001h]に対して実アドレスとして[0008h]が割り当てられているという変換情報が記述されている。

【0082】アドレス変換／アドレス管理処理部25はアドレス管理テーブルの更新完了後、書き込み要求、8KBのデータ、実アドレス[0008h]の情報をディスクアレイコントローラ14に送信する。

【0083】ディスクアレイコントローラ14は、受信したデータを分割し、分割したデータよりパリティデータを生成し、実アドレスと分割データ、パリティデー

15

タ、および書き込み要求をデバイスコントローラ15a~15eに送る。この場合、4つのディスク装置が接続されているのでディスクアレイコントローラはデータを4分割する。デバイスコントローラ15a~15dは受信した実アドレスに従って、分割データをディスク装置16a~16dに記録させ、デバイスコントローラ15eはパリティデータをディスク装置16eに記録させる。

【0084】以下、読み込み処理の場合について説明する。

【0085】ディスクアレイコントローラ14はホストアダプタ12、Alignment判定処理部13を介して送信されたホスト11の読み込み要求と論理アドレスを受信し、受信した論理アドレスをアドレス変換/アドレス管理処理部25が保有しているアドレス管理テーブルを参照して実アドレスに変換する。(図3のステップ310)。

【0086】(表3)にアドレス変換/アドレス管理処理部25が保有しているアドレス管理テーブルを示す。ディスクアレイコントローラ14は(表3)のアドレス管理テーブルより論理アドレス[0001h]に対応する実アドレス[0008h]を得、読み込み要求、実アドレス[0008h]をデバイスコントローラ15a~15eに送信し、デバイスコントローラ15a~15eは受信した実アドレスに従ってディスク装置16a~16eに記録されている分割データ、パリティデータを読み出し、ディスクアレイコントローラ14に送信する。ディスクアレイコントローラ14は受信したパリティデータでパリティチェックを行い、分割データを結合して、Alignment判定処理部13、ホストアダプタ12を介してホスト11に読み出したデータを送信する。

【0087】(実施の形態2)実施の形態2では、Alignment処理機能をディスクアレイコントローラ内部に持たせ、Misalignmentが発生した書き込み要求をAlignmentの書き込み要求にする場合について説明する。

【0088】図7は、本発明の第2の発明の一実施の形態のディスクアレイ装置の構成を示すブロック図である。

【0089】図7において、71はI/O処理を指示するホスト、72はホスト71と接続する装置を制御するホストアダプタ、73は書き込み時にはホストアダプタ72より入力されるデータにAlignment処理を行った後、データを分割し、分割したデータのパリティデータを生成し、読み込み時にはパリティチェックを行った後分割データを結合するディスクアレイコントローラ、74a~74eはディスクアレイコントローラ73より入力されたデータのI/O処理を制御するデバイスコントローラ、75a~75eはそれぞれデバイスコントローラ74a~74eの制御に従ってデータのI/O

16

処理を行うディスク装置である。

【0090】以上のように構成されたディスクアレイ装置について、図面を参照しながら説明する。

【0091】ホスト71は、ホストアダプタ72経由で、書き込み要求とデータの論理アドレス、およびデータをディスクアレイコントローラ73に送信する。ディスクアレイコントローラ73は受信したデータにAlignment処理、分割処理などを行う。

【0092】以下、ディスクアレイコントローラ73の処理を図8、図9を用いて詳細に説明する。図8はディスクアレイコントローラ73の詳細なブロック図、図9はディスクアレイコントローラ73の処理を示すフローチャートである。

【0093】図8において、81はホスト71から受信したコマンドを処理するコマンド処理部、82はコマンド処理部81から出力されたデータを受信し、データの書き込み処理に対してAlignment判定を行うAlignment判定処理部83へ送信するホストデータ受信/送信処理部、84はAlignment判定処理部83より出力されたデータを分割するホストデータ分割/結合処理部、85はAlignment判定処理部83より出力された判定結果に応じて分割データに無効データを補完する無効データ補完処理部、86はAlignment判定処理部83より出力された判定結果に応じて無効データ補完処理部85より出力されたデータの論理アドレスを実アドレスへ変換し、双方のアドレスの管理を行うアドレス変換/アドレス管理処理部、87はアドレス変換/管理処理部86より出力されたデータにパリティデータを生成するパリティ処理部である。

【0094】以下、書き込み処理の場合を説明する。

【0095】コマンド処理部81は、ホスト71から送信されてきたアクセス要求が書き込み処理であるか読み込み処理であるかの判定を行い、判定結果およびホスト71から受信したデータをホストデータ受信/送信処理部82を介してAlignment判定処理部83に送信する(図9のステップ901~ステップ902)。

【0096】Alignment判定処理部83は、受信した論理アドレスとデータのデータサイズからデータの書き込み処理がAlignmentであるか否かの判定を行い、判定結果を受信したデータと共にホストデータ分割/結合処理部84へ送信する(図9のステップ903)。

【0097】ここで、Alignment状態とは、以下の(1)(2)の条件を満たす場合を指す。

(1) データサイズブロックが、ハードディスクの最小単位1セクタ(512B)の整数倍であり、かつパリティ計算ブロック(512Bの整数倍サイズ)の整数倍である。

(2) データの書き込み開始位置を示すアドレスが、パリティ計算ブロックのAlignment境界であるオ

フセットと一致する。

【0098】また、2つの条件のいずれか一つでも満たさない状態をMisalignと定義する。

【0099】Alignment判定処理部83は受信したデータの書き込み処理がAlignmentであると判定した場合、受信したデータと判定結果をホストデータ分割/結合処理部84に送信する。

【0100】ホストデータ分割/結合処理部84は、受信したデータをパリティ計算ブロック単位に分割処理を施した後(図9のステップ904)、無効データ補完処理部85に判定結果と分割データを送信する。無効データ補完処理部85は受信した判定結果よりデータが条件(1)を満たしていると判断し、受信したデータをそのままアドレス変換/変換処理部86に送信する(図9のステップ905)。アドレス変換/管理処理部86は、受信した判定結果よりデータが条件(2)を満たしていると判定し(図9のステップ907)、受信したデータをパリティ処理部87に送信する(図9のステップ909)。判定結果がAlignmentである場合、図9のステップ909に示すアドレス変換は行なわれない。パリティ処理部87は受信したデータよりパリティデータを生成し(図9のステップ910)、書き込み要求、分割データ、パリティデータ、論理アドレスをデバイスコントローラ74a~74eへ送信する。

【0101】以下、Misalign状態の3つの場合における各処理部の処理を詳細に説明する。

【0102】まず、条件(1)のみを満たさないMisalign状態の場合について説明する。

【0103】この場合のデータの一例を図10の(a)の塗りつぶし部分に示す。図10において、処理単位となるパリティ計算ブロックサイズは4KB、パリティ計算ブロックにおける個々のブロックサイズは512Bである。図10の(a)のデータは、論理アドレス[0000h]、データサイズが6KBのデータである。

【0104】Alignment判定処理部83は受信した論理アドレスとデータサイズよりデータの書き込み処理がAlignmentであるかどうかを判定する(図9のステップ903)。Alignment判定処理部83は、論理アドレス[0000h]はパリティ計算ブロックのオフセット値を示すブロック番号[00h]と一致しているが、データサイズ6KBはパリティ計算ブロックサイズの4KBの倍数ではないことから、条件(1)のみ満たさないMisalignという判定結果Aと共にデータ、論理アドレス、書き込み要求をホストデータ分割/結合処理部84に送信する。

【0105】ホストデータ分割/結合処理部84は、受信したデータを4KBのパリティ計算ブロック単位に分割し、無効データ補完処理部85に、判定結果A、書き込み要求、論理アドレス、そして分割データを送信する(図9のステップ904)。この場合、データ用のディ

スク装置は4つ用意されているので、ホストデータ分割/結合処理部84はデータを4分割する。

【0106】無効データ補完処理部85は、ホストデータ分割/結合処理部84から送信されてきた判定結果Aより、条件(1)を満たさないMisalignであると判断し、図10の(a)に示す6KBのデータがパリティ計算ブロックの最小公倍数になるようになるように2KBの無効データを補完し、書き込み要求、無効データが追加された8KBのデータ、論理アドレス[0000h]、そして判定結果Aをアドレス変換/アドレス管理処理部86に送信する(図9のステップ905、906)。図10の(b)に無効データを追加された8KBのデータを示す。同図に示すように、図10の(b)のデータは図10の(a)のデータに斜線部に示す2KBの無効データを補完することで、パリティ計算ブロックの整数倍(2倍)のデータとなっている。

【0107】アドレス変換/アドレス管理処理部86は、受信した判定結果Aより、条件(2)を満たしていると判断し、直ちに、書き込み要求、無効データを補完した分割データ、論理アドレス[0000h]をパリティ処理部87に送信すると同時に、論理アドレス[0000h]より、実アドレスの割り当てを行い、保有していたアドレス管理テーブルの更新を行う(図9のステップ907、ステップ909)。(表4)にアドレス変換/アドレス管理処理部86が保有するアドレス管理テーブルを示す。

【0108】

【表6】

論理アドレス (ホストからのアドレス)	実アドレス (ディスクに書き込むブロック番号)
0000h	00h
000Ch	02h

【0109】(表6)に示すように、アドレス管理テーブルはデータの論理アドレスと実アドレスの対応した形で構成される。この場合、判定結果Aより受信したデータは条件(2)を満たしているため、論理アドレスと実アドレスは一致している。

【0110】アドレス変換/アドレス管理処理部86はアドレス管理テーブルの更新完了後、書き込み要求、分割データ、実アドレス[0000h]の情報をパリティ処理部87に送信する。

【0111】パリティ処理部87は受信したデータよりパリティデータを生成し、書き込み要求、データ、パリティデータ、および実アドレスをデバイスコントローラ74a~74eに送信する(図9のステップ910)。

【0112】次に、条件(2)のみを満たさないMisalignの場合について説明する。

【0113】この場合のデータの一例を図11の(a)の塗りつぶし部分に示す。図11において1ブロックは512Bのデータサイズブロックを示し、図11の(a)のデータは、処理単位となるデータサイズが4KB、論理アドレスが[0001h]のデータである。

【0114】Alignment判定処理部83は受信したデータを直ちに、ホストデータ分割/結合処理部84に送信するとともに、受信した論理アドレスとデータサイズより、データの書き込み処理がAlignmentであるかどうかを判定する(図9のステップ903)。Alignment判定処理部83は、データサイズ4KBはパリティ計算ブロックサイズが4KBの整数倍であるが、論理アドレス[0001h]とパリティ計算ブロックのオフセット値を示すブロック番号[00h]とが一致していないことから、条件(2)のみを満たさないMisalign状態という判定結果Bを、ホストデータ分割/結合処理部84に送信する。

【0115】ホストデータ分割/結合処理部84は、受信したデータを4KBのパリティ計算ブロック単位に分割し、無効データ補完処理部85に、判定結果A、書き込み要求、論理アドレス[0001h]、そして分割データを送信する(図9のステップ904)。この場合、データ用のディスク装置は4つ用意されているので、ホストデータ分割/結合処理部84はデータを4分割する。

【0116】無効データ補完処理部85は、ホストデータ分割/結合処理部84から送信されてきた判定結果Bより、受信したデータは条件(1)を満たしているため補完処理をする必要がないと判断し、書き込み要求、データ、論理アドレス[0001h]、および判定結果Bをアドレス変換/アドレス管理処理部86に送信する(図9のステップ905)。

【0117】アドレス変換/アドレス管理処理部86は、受信した判定結果Bより、受信したデータが条件(2)を満たしていないと判断し、受信した論理アドレスに対してアドレス変換を行う。アドレス変換/アドレス管理処理部86は、受信したデータの論理アドレス[0001h]を、書き込みを行おうとしていたパリティ計算ブロック[00h]の次のパリティ計算ブロックのオフセット値を示すブロック番号[01h]にシフトさせる為に、論理アドレス[0001h]に対して、7データサイズブロックだけシフトした実アドレス[0008h]の割り当てを行い、その変換情報をアドレス管理テーブルに保有する(図9のステップ907～ステップ909)。

【0118】(表7)にアドレス変換/アドレス管理処理部86が保有しているアドレス管理テーブルを示す。

【0119】

【表7】

論理アドレス (ホストからのアドレス)	実アドレス (ディスクに書き込むブロック番号)
0001h	01h
0009h	02h
	⋮
	⋮

【0120】(表7)に示すように、データの論理アドレス[0001h]に対応する実アドレスとして[0008h]が割り当てられているという変換情報が記述される。アドレス変換/アドレス管理処理部86はアドレス管理テーブルの更新完了後、書き込み要求、分割データ、実アドレス[0008h]の情報をパリティ処理部87に送信する。

【0121】パリティ処理部87は受信したデータよりパリティデータを生成し、書き込み要求、データ、パリティデータ、および実アドレスをデバイスコントローラ74a～74eに送信する(図9のステップ910)。

【0122】最後に、条件(1)、(2)を満たさないMisalign状態の場合について説明する。

【0123】この場合のデータの一例を図12の(a)の塗りつぶし部分に示す。図12において、処理単位となるパリティ計算ブロックサイズは4KB、パリティ計算ブロックにおける個々のブロックサイズは512Bである。図12の(a)のデータは、論理アドレス[0001h]、データサイズが4KBのデータである。

【0124】Alignment判定処理部83は受信したデータを直ちに、ホストデータ分割/結合処理部84に送信するとともに、受信した論理アドレスとデータサイズより、データの書き込み処理がAlignmentであるかどうかを判定する(図9のステップ903)。Alignment判定処理部83は、データサイズがパリティ計算ブロックサイズの4KBの倍数ではなく、論理アドレス[0001h]とパリティ計算ブロックのオフセット値を示すブロック番号[00h]と一致していないことから、条件(1)、(2)を満たさないMisalignであるという判定結果Cをホストデータ分割/結合処理部84に送信する。

【0125】ホストデータ分割/結合処理部84は、受信したデータを4KBのパリティ計算ブロック単位に分割し、無効データ補完処理部85に、判定結果A、書き込み要求、論理アドレス[0001h]、および分割データを送信する(図9のステップ904)。この場合、データ用のディスク装置は4つ用意されているので、ホストデータ分割/結合処理部84はデータを4分割する。

【0126】無効データ補完処理部85は、ホストデータ分割/結合処理部84から送信されてきた判定結果Cより、条件(1)を満たさないMisalign状態であると判断し、図12の(a)に示す6KBのデータがパリティ計算ブロックの最小公倍数になるようになるよ

うに2KBの無効データを補完し、無効データである事を示す処理を施した後に、書き込み要求、無効データが追加された分割データ、論理アドレス[0001h]、および判定結果Cをアドレス変換/アドレス管理処理部86に送信する(図9のステップ905、ステップ906)。図12の(b)に無効データを追加された8KBのデータを示す。同図に示すように、図12の(b)のデータは図6の(a)のデータデータに斜線部に示す無効データを補完することで、パリティ計算ブロックの整数倍(2倍)のデータとなっている。

【0127】アドレス変換/アドレス管理処理部86は、受信した判定結果Cより、受信したデータが条件(2)を満たしていないと判断し、受信した論理アドレスに対してアドレス変換を行う(図9のステップ907、ステップ908)。アドレス変換/アドレス管理処理部86は、受信したデータの論理アドレス[0001h]を、書き込みを行おうとしていたパリティ計算ブロック[00h]の次のパリティ計算ブロックのオフセット値を示すブロック番号[01h]にシフトさせる為に、論理アドレス[0001h]に対して、7データサイズブロックだけシフトした仮想アドレス[0008h]の割り当てを行い、その変換情報をアドレス管理テーブルに保有する(図9のステップ909)。

【0128】(表8)にアドレス変換/アドレス管理処理部86が保有しているアドレス管理テーブルを示す。

【0129】

【表8】

論理アドレス (ホストからのアドレス)	実アドレス (ディスクに書き込むブロック番号)
0001h	01h
000Dh	03h
	⋮
	⋮

【0130】(表8)に示すように、データの論理アドレス[0001h]に対して実アドレスとして[0008h]が割り当てられている変換情報が記述される。アドレス変換/アドレス管理処理部86はアドレス管理テーブルの更新完了後、書き込み要求、分割データ、実アドレス[0008h]の情報をパリティ処理部87に送信する。

【0131】パリティ処理部87は受信した分割データよりパリティデータを生成し、書き込み要求、分割データ、パリティデータ、および実アドレスをデバイスコントローラ74a~74eに送信する(図9のステップ910)。

【0132】デバイスコントローラ74a~74dは受信した実アドレスに従って、分割データをディスク装置74a~74dに記録させ、デバイスコントローラ74eはパリティデータをディスク装置74eに記録させ

る。

【0133】以下、読み込み処理の場合について説明する。

【0134】ディスクアレイコントローラ73はホストアダプタ72を介して送信されたホスト71の読み込み要求と論理アドレスを受信し、受信した論理アドレスをアドレス変換/アドレス管理処理部86が保有しているアドレス管理テーブルを参照して実アドレスに変換する。(図9のステップ911)。

10 【0135】(表8)にアドレス変換/アドレス管理処理部86が保有しているアドレス管理テーブルを示す。ディスクアレイコントローラ73は(表8)のアドレス管理テーブルより論理アドレス[0001h]に対応する実アドレス[0008h]を得、読み込み要求、実アドレス[0008h]をデバイスコントローラ74a~74eに送信し、デバイスコントローラ74a~74eは受信した実アドレスに従ってディスク装置75a~75eに記録されている分割データ、パリティデータを読み出し、ディスクアレイコントローラ73に送信する。ディスクアレイコントローラ73は受信したパリティデータでパリティチェックを行い、分割データを結合して、ホストアダプタ72を介してホスト71に読み出したデータを送信する。

【0136】なお、本発明において、条件(1)を満たさないMisalignである場合に、無効データを補完せずに、アドレス変換/アドレス管理部で次のデータの書き込み開始位置を、前回書き込まれたデータの最終パリティ計算ブロックの次のパリティ計算ブロックのオフセットに一致させることにより、動作的には、常にAlignment状態と同じ様な処理(擬似Alignment)が実現できる。

30 【0137】また、本発明において、条件(1)と(2)の双方を満たさないMisalignである場合に、アドレス変換/管理処理部は、初めて記録するデータは、データの書き込み開始位置を示すアドレスが、パリティ計算ブロックのオフセットと一致していないので、データを書き込もうとしていたパリティ計算ブロックの次のパリティ計算ブロックのオフセットに、データの書き込み開始位置を変更し、次のデータの書き込み開始位置を、前回書き込まれたデータの最終パリティ計算ブロックの次のパリティ計算ブロックのオフセットに一致させることにより、動作的には、常にAlignment状態と同じ様な処理(擬似Alignment)が実現できる。

【0138】また、本発明では、Alignment処理を行うために無効データ補完処理部とアドレス変換/管理処理部を具備したが、条件(1)、または(2)のどちらかを必ず満たす前提条件がある場合は、必ずしも双方の処理を具備しなくてもよい。

50 【0139】また、本発明では、アドレスの変換の方法

として、アドレス管理を行うアドレス管理テーブルを設けた場合について述べたが、固定長のアドレスをシフトさせる場合、レジスタなどを用いて、アドレス変換を実施した場合でも同様の効果が得られる。

【0140】また、本発明では、処理単位となるパリティ計算ブロックサイズを4KBとして説明したが、これに限定するものではなく、512Bの整数倍サイズであればよい。

【0141】また、本発明では、アドレス管理テーブルには論理アドレスとそれに対応する実アドレスを記述したが、これに限定するものではない。

【0142】

【発明の効果】以上の説明から明かなように、本発明のディスクアレイ装置によれば、Misalignが発生する書き込み要求があった場合、後続のデータの書き込みをAlignmentで行えるようデータの補完やアドレス変換を行うAlignment処理機能を有することで、後続のI/O処理をAlignmentで行うことができ、Alignment状態と同様の性能を発揮できる。

【図面の簡単な説明】

【図1】本発明の実施の形態1におけるディスクアレイ装置の構成を示すブロック図

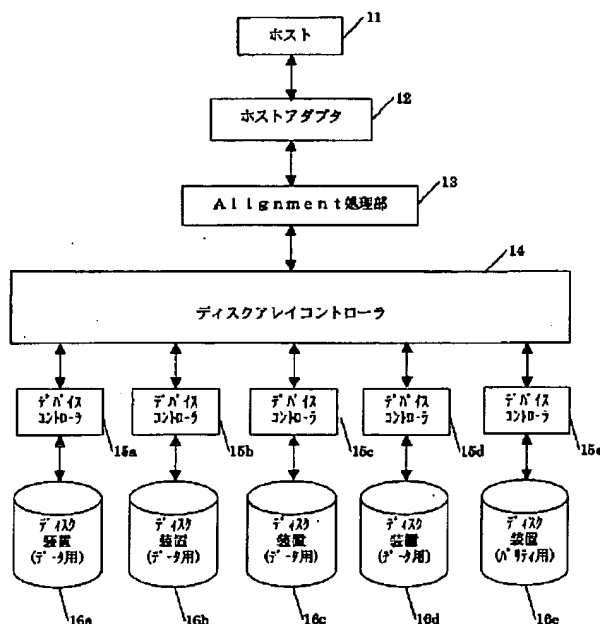
【図2】同Alignment処理部の詳細な構成を示すブロック図

【図3】同Alignment処理部の動作を説明するフローチャート

【図4】同無効データ補完処理の一例を説明する図

【図5】同アドレス変換処理の一例を説明する図

【図1】



【図6】同無効データ補完処理およびアドレス変換処理の一例を説明する図

【図7】本発明の実施の形態2におけるディスクアレイ装置の構成を示すブロック図

【図8】同Alignment処理部の詳細な構成を示すブロック図

【図9】同Alignment処理部の動作を説明するフローチャート

【図10】同無効データ補完処理の一例を説明する図

【図11】同アドレス変換処理の一例を説明する図

【図12】同無効データ補完処理およびアドレス変換処理の一例を説明する図

【図13】従来のディスクアレイ装置の構成を示すブロック図

【図14】従来のディスクアレイコントローラの詳細な構成を示すブロック図

【図15】従来のディスクアレイコントローラの動作を説明するフローチャート

【図16】従来のディスクアレイ装置の書き込み処理を説明する図

【符号の説明】

11、71 ホスト

12、72 ホストアダプタ

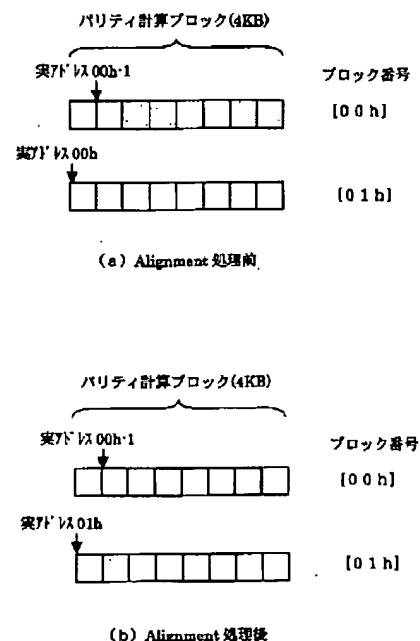
13 Alignment処理部

14 73 ディスクアレイコントローラ

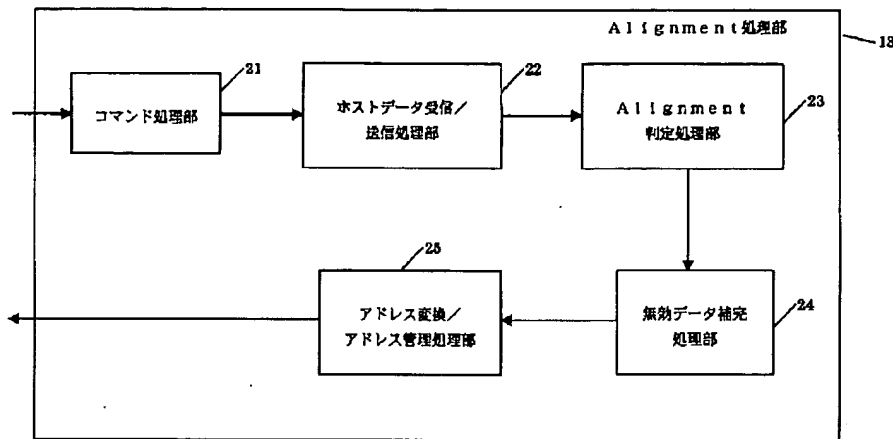
15a~15e、74a~74e デバイスコントローラ

16a~16e、75a~75e ディスク装置

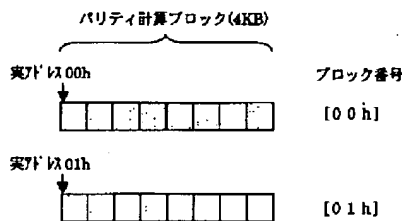
【図5】



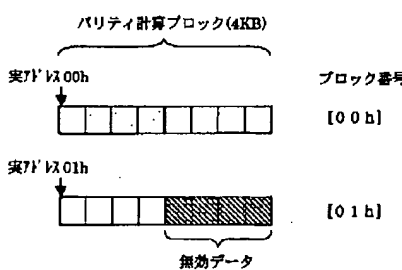
【図2】



【図4】

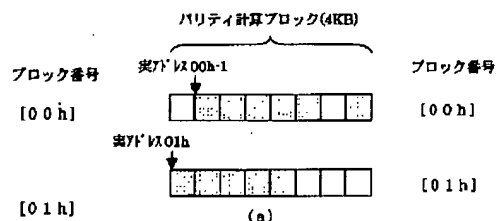


(a) Alignment 処理前

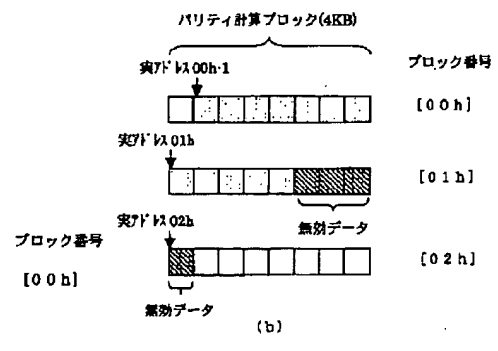


(b) Alignment 処理後

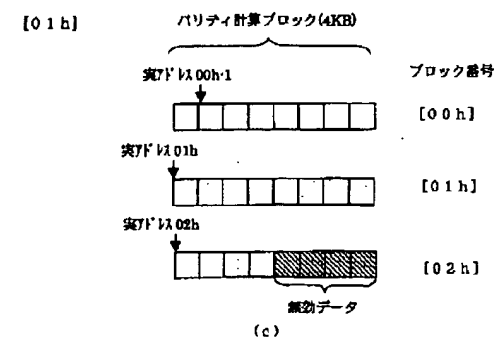
【図6】



(a)

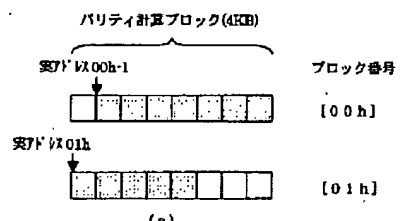


(b)

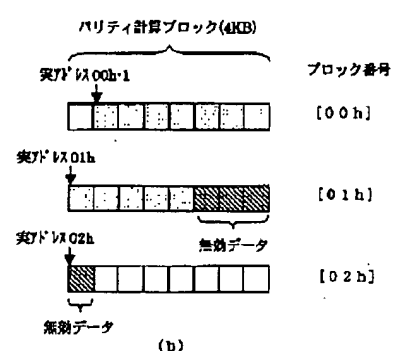


(c)

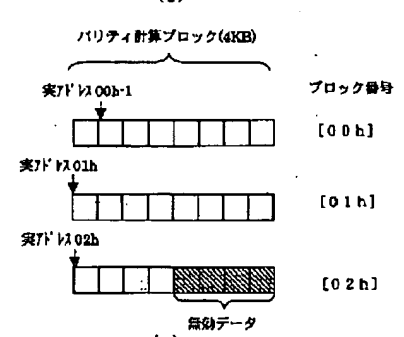
【図12】



(a)

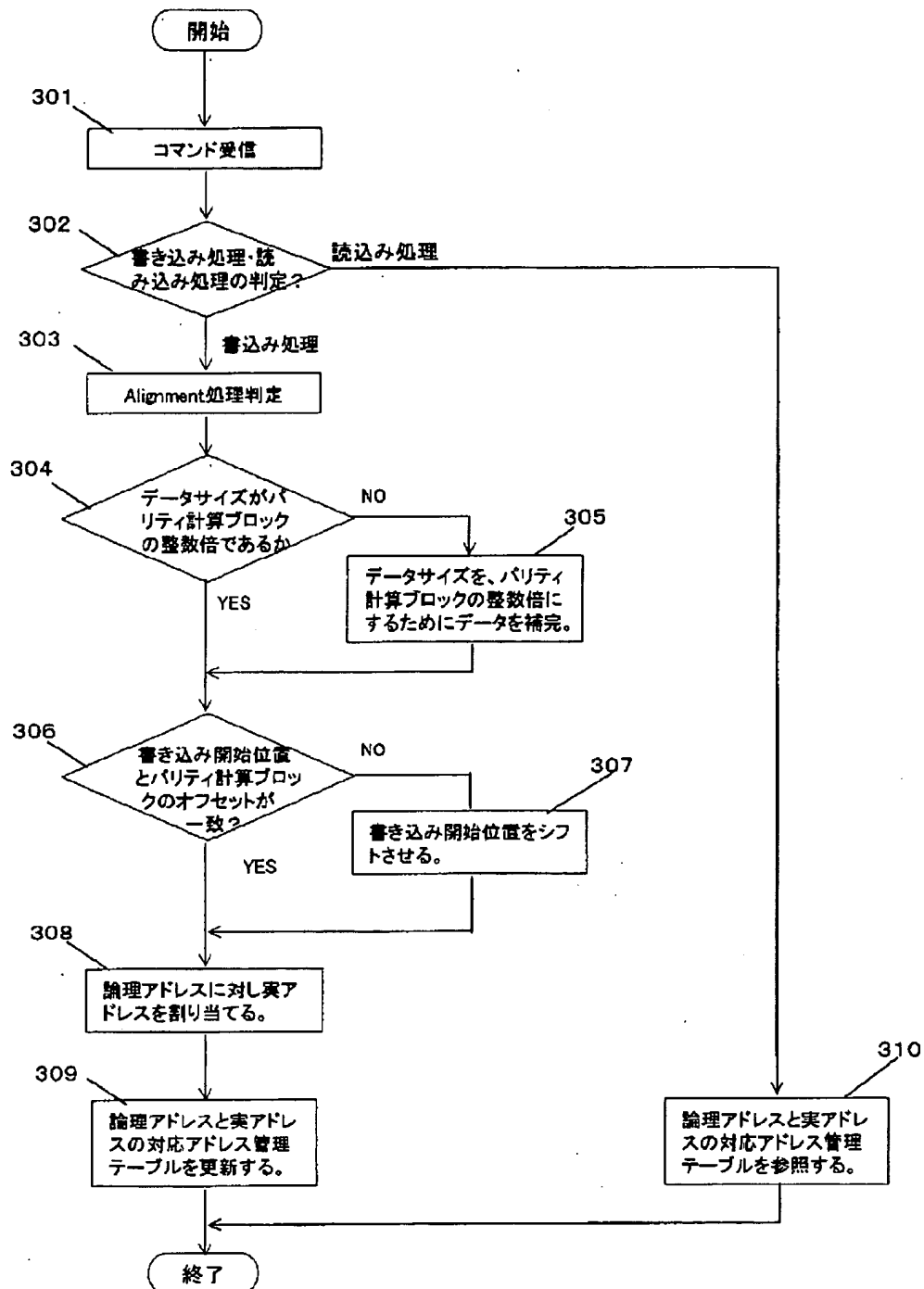


(b)

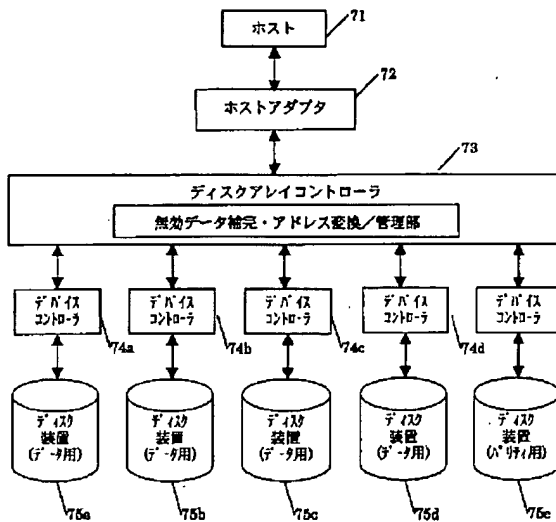


(c)

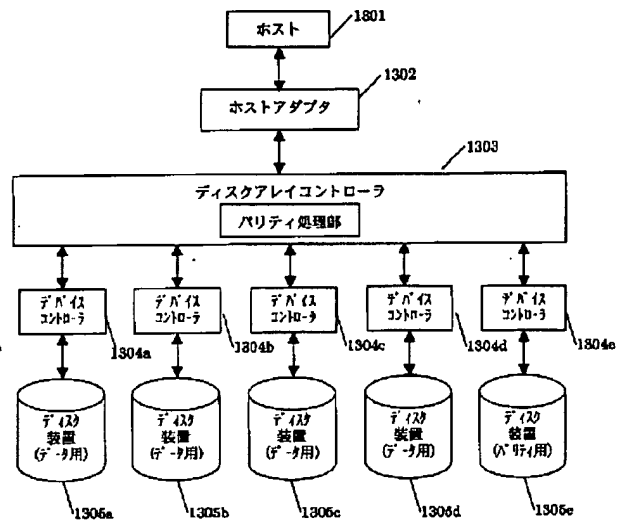
【図3】



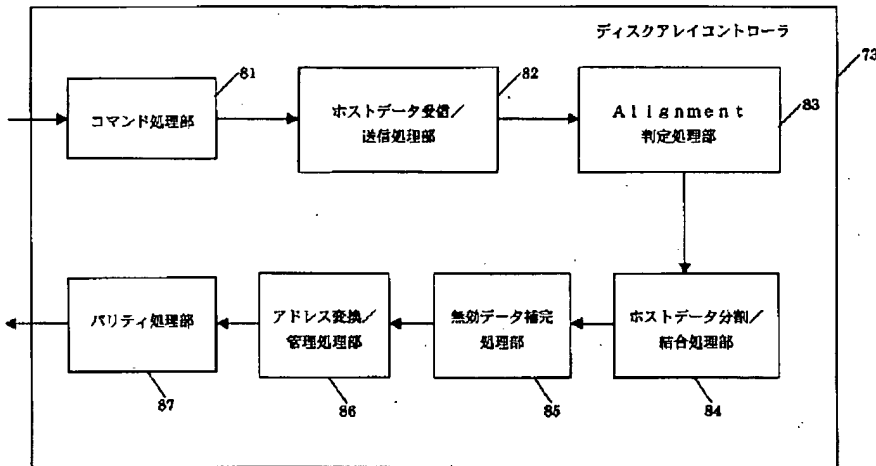
【図7】



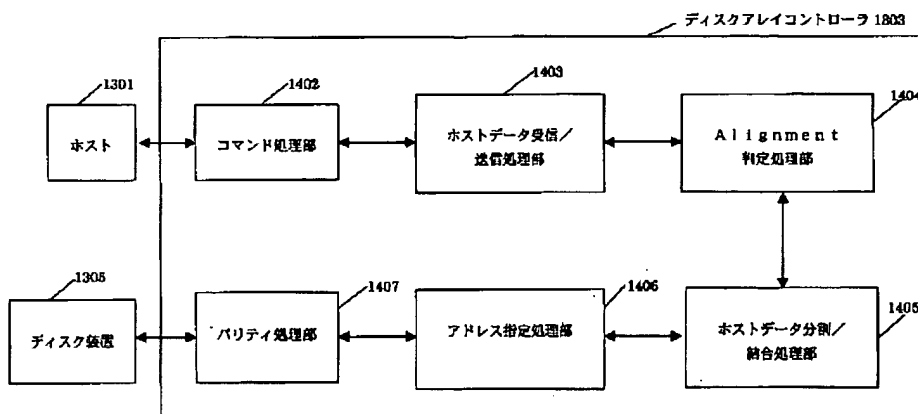
【図13】



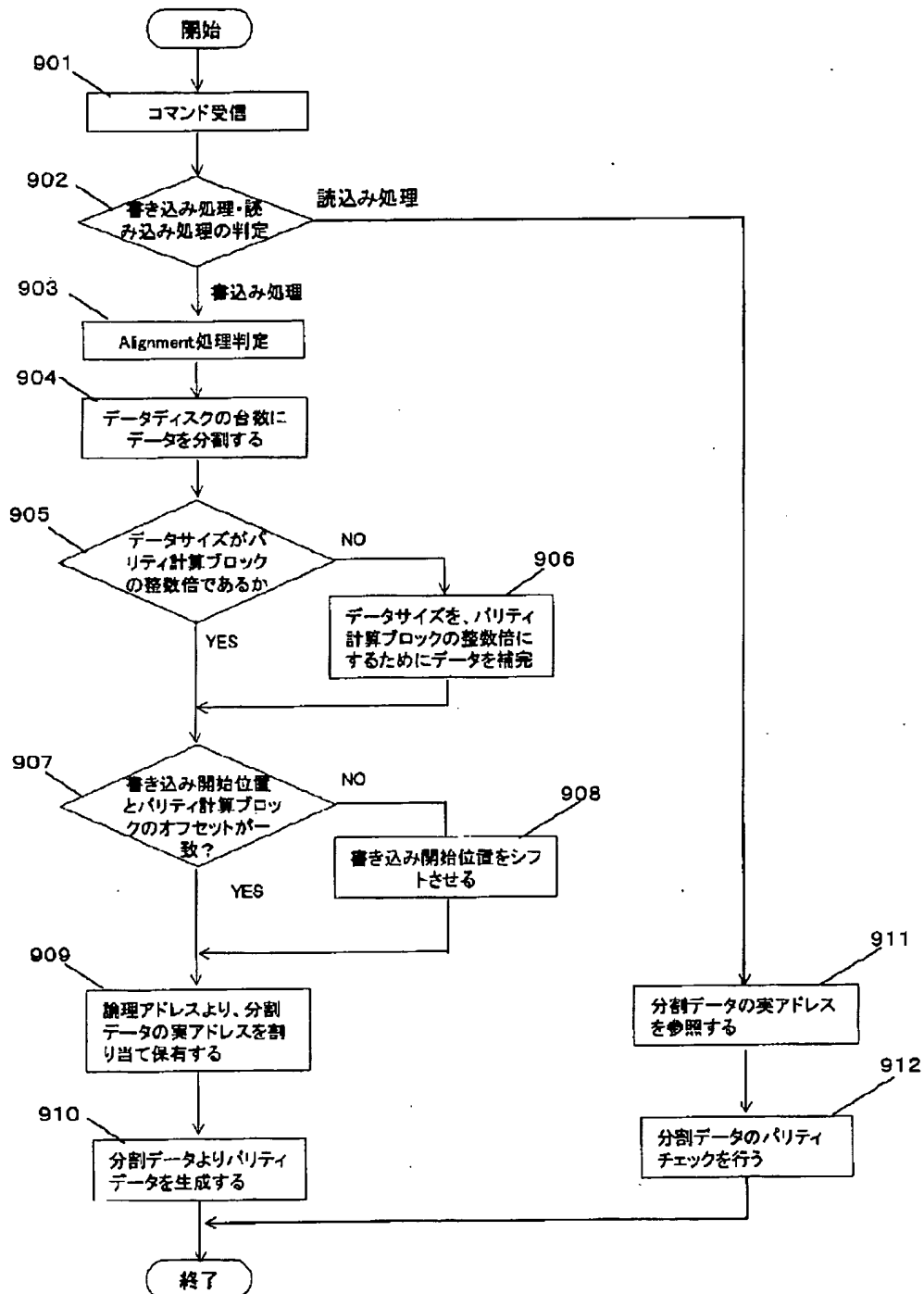
【図8】



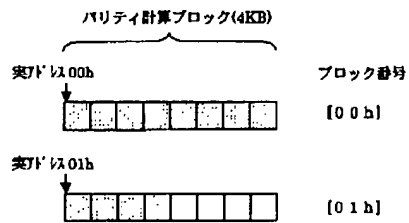
【図14】



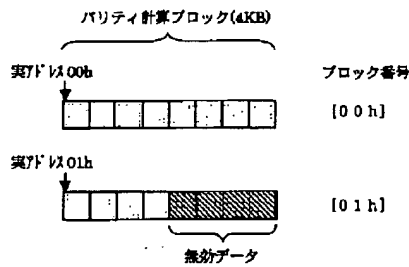
【図9】



【図10】

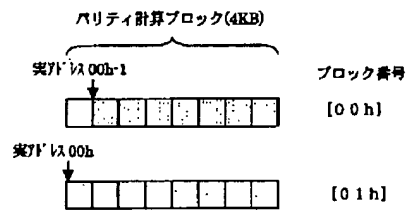


(a) Alignment 処理前

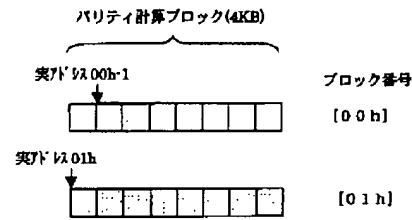


(b) Alignment 処理後

【図11】

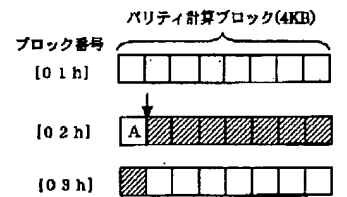
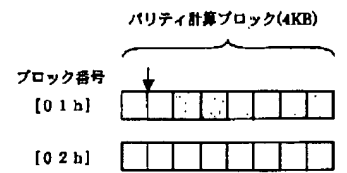
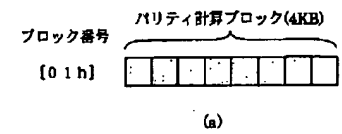


(a) Alignment 処理前



(b) Alignment 処理後

【図16】



【図15】

